DNA Binding Induces a Nanomechanical Switch in the RRM1 Domain of TDP-43

Yong Jian Wang[†,#], *Palma Rico-Lastres*[†,#], *Ainhoa Lezamiz*[†,#], *Marc Mora*[†], *Carles Solsona*[‡], *Guillaume Stirnemann*[&] *and Sergi Garcia-Manyes*[†,*]

[†] Department of Physics and Randall Centre for Cell and Molecular Biophysics, King's College London, WC2R 2LS, London, UK.

[‡] Department of Pathology and Experimental Therapeutics, Faculty of Medicine and Health Sciences, University of Barcelona and Bellvitge Biomedical Research Institute (IDIBELL) L'Hospitalet de Llobregat, Barcelona 08907, Spain

[&] CNRS Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, Univ. Paris Denis Diderot, Sorbonne Paris Cité, PSL Research University, 75005 Paris, France

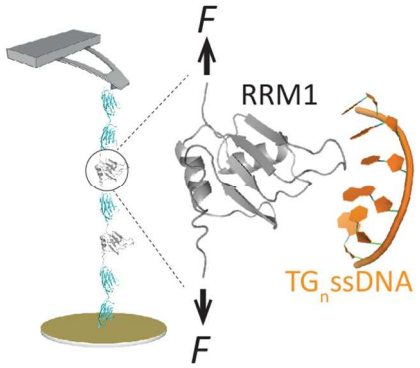
# These authors contributed equally to the work

Corresponding Author

* sergi.garcia-manyes@kcl.ac.uk.

ABSTRACT. Understanding the molecular mechanisms governing protein-nucleic acid interactions is fundamental to many nuclear processes. However, how nucleic acid binding affects the conformation and dynamics of the substrate protein remains poorly understood. Here we use a combination of single molecule force spectroscopy AFM and biochemical assays to show that the binding of TG-rich ssDNA triggers a mechanical switch in the RRM1 domain of TDP-43, toggling between an entropic spring devoid of mechanical stability and a shock absorber bound-form that resists unfolding forces of ~40 pN. The fraction of mechanically resistant proteins correlates with an increasing length of the TG_n oligonucleotide, demonstrating that protein mechanical stability is a direct reporter of nucleic acid binding. Steered Molecular Dynamics simulations on related RNA oligonucleotides reveal that the increased mechanical stability fingerprinting the holo-form is likely to stem from a unique scenario whereby the nucleic acid acts as a "mechanical staple" that protects RRM1 from mechanical unfolding. Our approach highlights nucleic acid binding as an effective strategy to control protein nanomechanics.

TOC GRAPHICS



KEYWORDS. Single-molecule studies, RRM1 domain, Nucleotides, Mechanical properties, Nucleic acid binding.

The nanomechanical properties of individual proteins regulate a number of major biological processes, including the deformability of the extracellular matrix[1], mechanotrans duction in focal adhesion adaptors[2], the elasticity of cardiac titin[3] or the degradation of proteins in the proteasome[4-5]. Several molecular tactics have been convincingly used to modify the mechanical stability of proteins; beyond simple protein unfolding —which converts the mechanically resistant native state into a compliant and extended protein conformation[6]—, the introduction of key point mutations in the so-called "mechanical clamp"[7] and the presence of stiff disulfide bonds[8-9] or covalent organometallic bonds [10-11] have all been shown to have a direct effect on the mechanical stability of the natively folded conformation.

In addition to these more common strategies, ligand binding[12] has recently emerged as an effective orthogonal modulator of protein nanomechanics. For example, DHFR was shown to increase its mechanical stability upon binding nicotinamide adenine dihydrogen phosphate (NADPH), 7,8-dihydrofolate (DHF) or inhibitor methotrexate (MTX), converting a purely elastic protein devoid of mechanical stability into an efficient shock absorber able to withstand stretching forces[13]. Likewise, the enzyme staphylococcal nuclease increases its mechanical resistance upon binding its inhibitor deoxythymidine 3',5'- bisphosphate[14]. Similarly, binding of small sugars (in the case of maltose binding protein[15-16], the hyperthermophilic adenine diphosphate (ADP)-dependent glucokinase[17] and membrane transporters[18]) or single amino acids (such as leucine[19]) can change both the height of the energy barriers and the distribution of

unfolding pathways. In addition, metal binding has also revealed as a successful strategy to regulate protein stiffness through calcium [20-23] or nickel [24-25] binding. Perhaps even more conspicuous are the mechanical consequences of small peptide —or full protein— binding. In this vein, the mechanical properties of protein G are substantially increased upon binding the IgG antibody[26]. Furthermore, SUMO1 increases its mechanical stability upon binding small peptides[27]. Similarly, the attachment of the short hydrophobic APPY polypeptide induces selective increase of the mechanical properties of the domain I of the multidomain DnaJ chaperone[28]. Other recent examples epitomise the importance of mechanically revealing key binding pockets that are otherwise hidden in the folded conformation [29 30-32].

Collectively, these experiments revealed the large knock-on effects that protein-protein interactions have on protein nanomechanics. This growing experimental evidence makes it tempting to speculate that, given the increasingly large number of identified DNA- and RNA-binding proteins (DRBPs)[33], nucleic acid binding, further to modifying unfolding pathways[34], might play an analogous role in regulating the mechanical stability of proteins. Direct testing of this hypothesis has remained elusive, mostly due to the lack of an extensive pool of DNA-binding proteins for which the crystal structure in the apo- and holo- forms has been solved, and especially given the difficulty to obtain these proteins biochemically free of the nucleic acid partner. Here, we investigated how nucleic acids of well-defined sequences regulate the nanomechanical properties of the RRM1 domain of the 43 kDa TAR DNA-binding protein (TDP-43). TDP-43 plays important roles in many essential cellular functions involved in DNA transcription and RNA translation[35], and it is hence capable of binding both RNA and DNA. Furthermore, TDP-43 has been associated to several important neurodegenerative disorders,

including amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD)[36-38]. Under physiological conditions, TDP-43 is predominantly localised in the nucleus with low levels in the cytoplasm[39-40] and conducts a multiplex of functionalities, being involved in different steps of RNA processing including transcription, mRNA splicing, transport and translation, and works as a transcription factor as well[41]. Given its multifunctional role, accessing the microscopic insights underpinning the protein-nucleic acid interaction is of capital importance towards establishing a link between function and protein conformation. From the topological perspective, TDP-43 is composed of two tandem RNA recognition motifs (RRM1 and RRM2) flanked between an N-terminal domain (NTD), an NLC segment that has been reported to bind RNA[42], and the C-terminal glycine-rich domain (GRD, Figure 1a)[43]. The crystal structure of both RRM domains has been solved in complex to different UG- and TG- rich single-stranded RNA and DNA sequences[44], concluding that RRM1 plays a dominant role in nucleic acid binding whereas RRM2 holds a supporting function[45]. Given its ability to effectively bind DNA and RNA, and thanks to the fact that its binding properties have been characterised both structurally and biochemically, RRM1 emerges as an excellent case study to elucidate how nucleic acid binding has direct effects on protein conformation.

Using single molecule force spectroscopy in combination with biochemistry assays, here we demonstrate that DNA binding has a major effect on the mechanical properties of the RRM1 domain of TDP-43, triggering its transition from a compliant, entropic spring into a mechanically resistant shock absorber. Supporting Molecular Dynamics simulations reveal that such a mechanical switch results from a unique molecular strategy whereby the nucleic acid functions as a 'molecular lid' that protects the RRM1 domain from mechanical unfolding.

To investigate the mechanical properties of the apo-RRM1 domain, we constructed a polyprotein containing two RRM1 monomers, each one flanked by three titin I27th Ig domains that serve as standard molecular fingerprints, resulting in the [$I27_2$-RRM1-I27-RRM1-$I27_2$] polyprotein (Figure 1a). A multistep elution protocol ensured quantitative removal of DNA[46], confirmed by the low (~0.6) ratio of absorbance measured at 260/280 nm, a generally accepted signature of quantitative removal of DNA[46]. Individual polyproteins were stretched under a constant velocity of 400 nm s^{-1} using an atomic force spectrometer (AFM). The resulting force-extension trajectories exhibited a first feature-less protein extension (associated to a total contour length, L_T = 81 nm) followed by the well-characterised unfolding of the I27 protein, occurring at forces ~200 pN and hallmarked by ΔL_{I27} = 28 nm[47] (Figure 1c). Having learnt that apo-RRM1 is void of mechanical stability, we sought to examine whether the addition of TG_{15}, which effectively binds RRM1 as revealed by electrophoretic mobility shift assays (EMSA, Figure 1b), has an effect on the nanomechanical properties of RRM1. The individual unfolding trajectories (Figure 1d) show that, in sharp contrast to the apo-form, mechanical unfolding of the holo-RRM1 results in a well-defined force peak occurring at forces F = 39 ± 11 pN that occurs concomitant with an increment in contour length of ΔL = 28 ± 1 nm (Figure S1), which is consistent with the complete extension of the RRM1 domain (77aa x 0.38[6] nm/aa – folded length(1.3nm) =27.96 nm) after mechanical unfolding. A gallery of individual unfolding trajectories for both the apo- and holo- proteins forms is shown in the Figure S2. To further confirm that the unfolding events observed in Figure 1d corresponds to the DNA-mediated mechanical stabilisation of RRM1, we repeated the same experiments in the presence of CA_{15}, which does not quantitatively bind RRM1 (Figure 1b)[48-49]. Under these control conditions, the

individual unfolding trajectories lacked the well-defined force peak (Figure 1e), hence mostly recapitulating the behaviour of apo-RRM1.

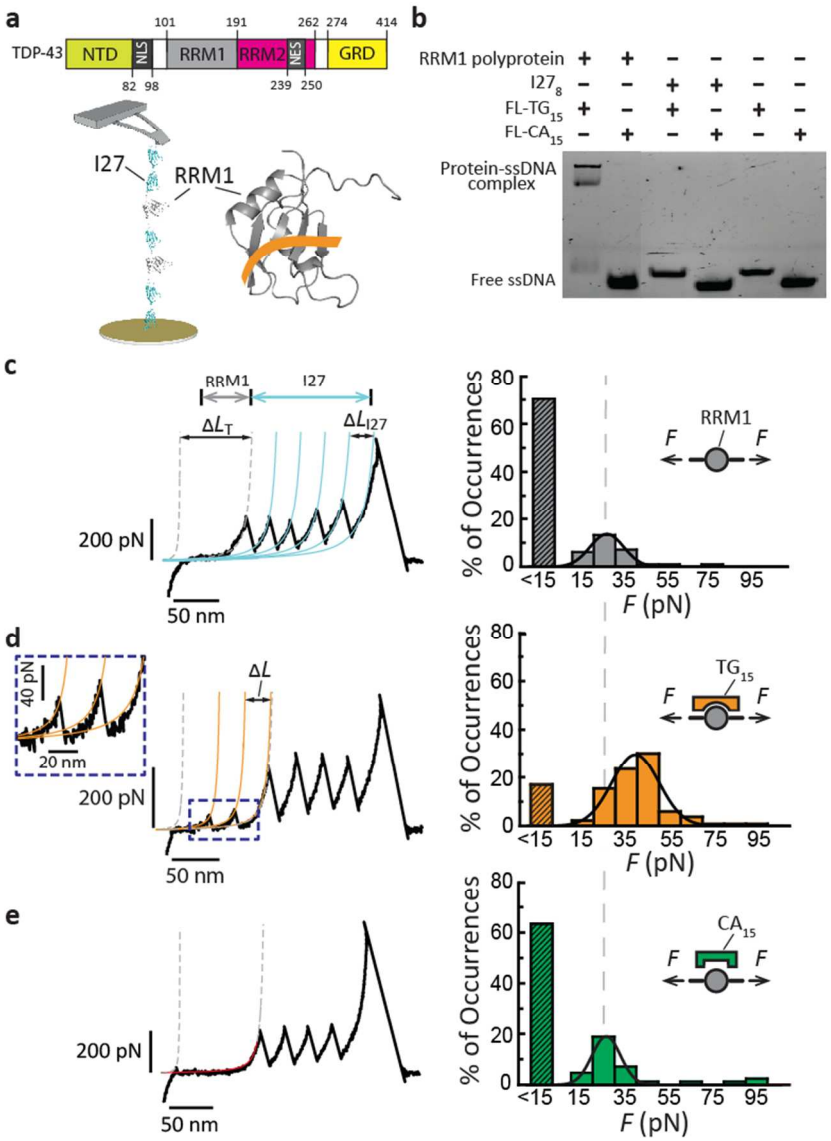


Figure. 1. (a) Schematic representation of the TDP-43 protein, composed of two non-equivalent RRM domains. A polyprotein containing two RRM1 (grey) domains, [I27$_2$-RRM1-I27-RRM1-I27$_2$], is stretched between a gold surface and an AFM cantilever tip, and its nanomechanical

properties were tested upon oligonucleotide binding (orange, PDB: 4BS2). (b) DNA specifity assay monitored by EMSA performed on RRM1 and $I27_8$ polyprotein constructs upon addition of different FL ssDNA oligonucleotides. The amount of DNA and protein was maintained fixed (10 nM and 80 nM, respectively, in a 14 μl load). The binding assay was performed in a 25 mM HEPES pH 7.4 buffer solution. (c) *Left*: Individual unfolding trajectory of the [$I27_2$-RRM1-I27-RRM1-$I27_2$] polyprotein, exhibiting a first feature-less extension corresponding to the unfolding of the RRM1 domains followed by the unfolding of the I27 markers, occurring at forces ~200 pN. *Right*: Histogram of the forces required to unfold the apo-RRM1 domain, demonstrating that in ~70% of the unfolding events (n = 69) the apo-RRM1 form unfolds in the absence of mechanical stability. (d) Upon addition of ~1.5 μM of $ssTG_{15}$ oligonucleotide, the mechanical unfolding of the holo-RRM1 domain can be fingerprinted by a saw-tooth pattern, with the unfolding peaks occurring at forces ~39 pN and concomitant to an increment in contour length of ΔL ~28 nm, n = 113. (e) Analogous control experiments using a CA_{15} oligonucleotide (n = 54), which do not quantitatively bind RRM1, exhibit unfolding trajectories devoid of mechanical stability that recapitulate the apo-form.

Inspired by recent biochemistry findings[49-50], we then asked whether the length of (TG)-containing ssDNA oligonucleotides has an effect on RRM1 binding in our experimental conditions. To address this question, electrophoretic mobility shift assays were used to characterise the binding of fluorescently-labelled (henceforth, FL) TG_3, TG_4, TG_5, TG_6, TG_{15} oligonucleotides (10 nM) with increasing stoichiometric ratios of the [$I27_2$-RRM1-I27-RRM1-$I27_2$] polyprotein (Figure 2a). These experiments demonstrated that longer TG oligonucleotides require a lower protein:DNA ratio to induce quantitative binding (Figure 2a and

Figure S3). To test whether the measured binding efficiency directly correlates with the DNA-induced mechanical stabilization of RRM1, we examined the mechanical behaviour of RRM1 when exposed to different TG-containing ssDNA oligomers of varying lengths. Crucially, in all cases we observed signatures of mechanical stabilization of RRM1 (Figure 2b, Figure S4). Remarkably, the fraction of trajectories displaying a mechanical peak increased with the number of TG repetitions, with a significant transition towards the bound fraction for TG_n>5 (Figure 2c), thus recapitulating the binding titration measured in Figure 2a. Altogether, these experiments suggest that the change in mechanical stability of RRM1 is a direct read-out of DNA binding.

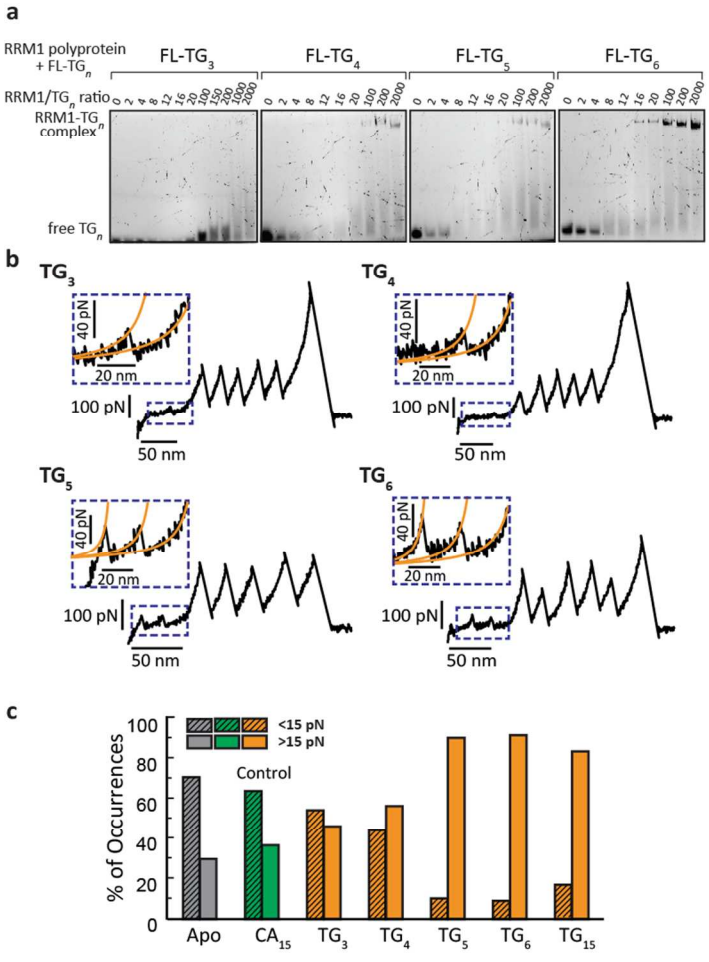


Figure 2. (a) Protein:DNA titrations monitored by EMSA. FL TG_3, TG_4, TG_5 and TG_6 ssDNA oligonucelotides (10 nM) titrated with increasing amounts of the RRM1 polyprotein construct in 25 mM HEPES pH 7.4 buffer. Values on the top of the gel indicate the RRM1:DNA ratio. (b) Individual unfolding trajectories corresponding to the unfolding of RRM1 domains in the presence of TG_3, TG_4, TG_5, and TG_6 single stranded oligonucleotides, exhibiting force peaks corresponding to the forced unfolding. (c) Histogram corresponding to the % of individual unfolding events (n = 40-140) featuring mechanical stability under all tested conditions, with the % increasing with the length of the TG_n oligonucleotide.

We next set out to probe whether the recovery of mechanical stability through DNA binding is a reversible process intricately linked to mechanical refolding. To this purpose, we conducted force-quench experiments (which afford superior control of the folding dynamics)[51] in the presence of TG_{15}, whereby the force was first ramped up to 240 pN at a constant rate of 40pN s^{-1} to trigger the unfolding of holo-RRM1 (fingerprinted by a step-wise increase of protein length of 19 ± 1 nm, occurring at F = 30 ± 8 pN, Figure S5), followed by the unfolding of the I27 domains, marked by the increase of the protein's contour length in ~25 nm steps (Figure 3a). The force was subsequently withdrawn for t_q = 15 s to trigger protein refolding before the force was ramped up again to test the folding status of the protein. Remarkably, in ~46 % of the trajectories (n = 13), mechanically re-stretching of the protein mirrored the initial unfolding sequence whereby the mechanically resistant RRM1 was first unfolded, prior to the unfolding of the I27 domains occurring at higher forces. The recovery of mechanical stability for RRM1 suggests that, upon refolding, RRM1 is able to effectively re-bind TG_{15}. Similar conclusions were

qualitatively reached in the case of TG_6 binding (Figure S6). Two distinct scenarios could mechanistically explain the rebinding process; either DNA was not removed from the protein and kept bound after unfolding, or alternatively DNA was able to re-bind from the solution within the experimental quench time (Figure 3b). Discriminating between both possibilities would eventually require an extensive set of experiments where both the quench time and also the time the mechanically unfolded protein is exposed to the solvent environment are independently varied. Repeating the experiments over a range of protein concentrations could also help elucidate whether or not nucleic acids remain bound after RRM1 mechanical unfolding. While these experiments are beyond the scope of the present work, our results highlight the reversibility of the binding process on experimental timescales, fingerprinted by the recovery of the RRM1 mechanical stability.

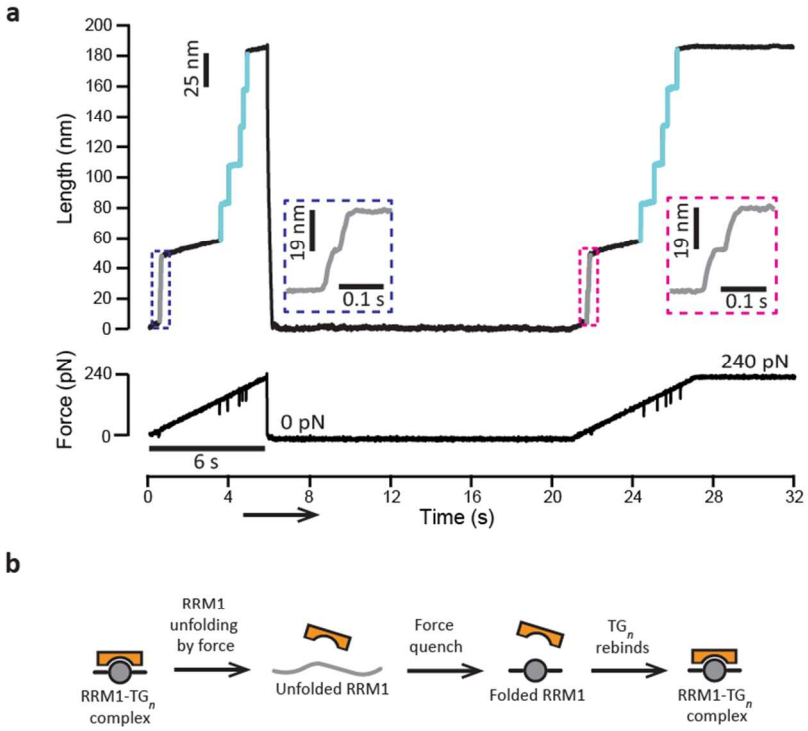


Figure 3. (a) Individual refolding trajectory (n = 13) of a single [$I27_2$-RRM1-I27-RRM1-$I27_2$] polyprotein in the presence of TG_{15} following a force-quench protocol. During the first 6 seconds, the protein was stretched at a constant loading rate of 40 pN s^{-1}. This first force-ramp pulse elicited the initial unfolding of the RRM1 domain, fingerprinted by the presence of steps of 19 ± 1 nm (grey) and occurring at forces 30 ± 8 pN. The mechanical unfolding of the I27 marker occurred at higher forces ~150-210 pN concomitant to a protein 25 nm-stepwise length increase (blue). Once the protein reached a stretching force of 240 pN, the force was completely removed for t_q = 15 s to trigger protein refolding. The *test* pulse, mirroring the first force-ramp protocol, probes the folding status of the protein. Upon re-stretching the protein at a constant rate of 40 pN s^{-1}, RRM1 re-unfolded as hallmarked by the presence of 19 nm steps, followed by the re-unfolding of the I27 domains. The complete recovery of mechanical stability of the RRM1 domain is an unambiguous proxy for the successful (*i*) refolding of the protein and the subsequent rebinding of the TG_{15} ssDNA oligonucleotide. (b) Two possible scenarios could account for the unbinding and rebinding of TG_{15} under force. In the first scenario, upon protein unfolding, the ssDNA would be completely removed. In this case, rebinding a ssDNA molecule from solution would occur during the quench time. Alternatively, after mechanical unfolding the TG_{15} oligonucleotide might remain attached to RRM1, facilitating the conformational search for re-binding upon removal of the pulling force.

To obtain an atomistic picture of the mechanism by which DNA enhances the mechanical unfolding of RRM1, we conducted Steered Molecular Dynamics (SMD) simulations under constant force conditions (Figure 4). Lacking the structure of the RRM1 bound to the DNA TG sequence, we used instead as a proxy the RRM1 structure bound to the RNA sequence

(5'-GUGUGAAU-3') for which the solution NMR structures are available (PDB:4BS2)[44] . We performed 15 independent simulations for each protein state (holo- or apo-) and pulling scenarios (N-terminal or C-terminal pulling). Comparing the mean unfolding time of the apo-form when stretched at a constant-force of F = 160 pN (grey trajectories) with that measured for the RNA-bound protein (orange trajectories) displays a significant separation of timescales (Figure 4a,b), suggesting that RNA binding slows down the mechanical unfolding of RRM1. In both cases, pulling from the N-terminus leads to unfolding times that are noticeably larger than those observed when the protein is pulled from the C-terminus (Figure 4a,b), implying that pulling from both termini is not a fully-equivalent process, in the sense that the protein-nucleic acid tandem (Figure 4c) resists better mechanical stress when the force propagates along the N- to the C- termini direction (Figure 4d). Close investigation to the dynamics of the unfolding trajectories put forward a rather unique mechanism of nucleic acid-mediated mechanical stabilization of RRM1; in the apo- form, mechanical unfolding occurs upon disruption of the hydrogen bonds present between ß1 and ß5. The protein geometry is such that the ß1-ß5 strands are not aligned with the direction of force application, resulting in a low cooperativity of the hydrogen-bonds and thus very low mechanical resistance (unfolding occurs within ~10 ns in our simulations at F = 160 pN, a force at which well-studied mechanically-resistant proteins such as the I27 markers would not unfold in the simulations). In the native structure of the holo-RRM1, RNA binds directly on top of ß1 and ß5, the protein region from where unfolding starts (Figure 4c,e). Detailed scrutiny of the individual trajectories (Supplementary video1) shows that RNA acts as a "molecular staple" that sits on top of the stretched RRM1, preventing its unfolding (Figure 4e, *left*). Thermal fluctuations coupled with the mechanical stress applied to the protein eventually displace RNA from this 'lid' well defined-position (Figure 4e, *middle*). Only after RNA has

been removed can the protein proceed to rapidly unfold and extend (Figure 4e, *right*). Hence, RNA functions as an effective molecular 'stopper' that delays protein unfolding, occurring only after RNA has been displaced from its binding position. As a further confirmation of this mechanism, we performed 5 additional simulations where RNA was maintained fixed (Fig. S7). No unfolding was observed in that case, whereas close to 100% of the traces in the presence of flexible RNA unfolded on that same timescale.

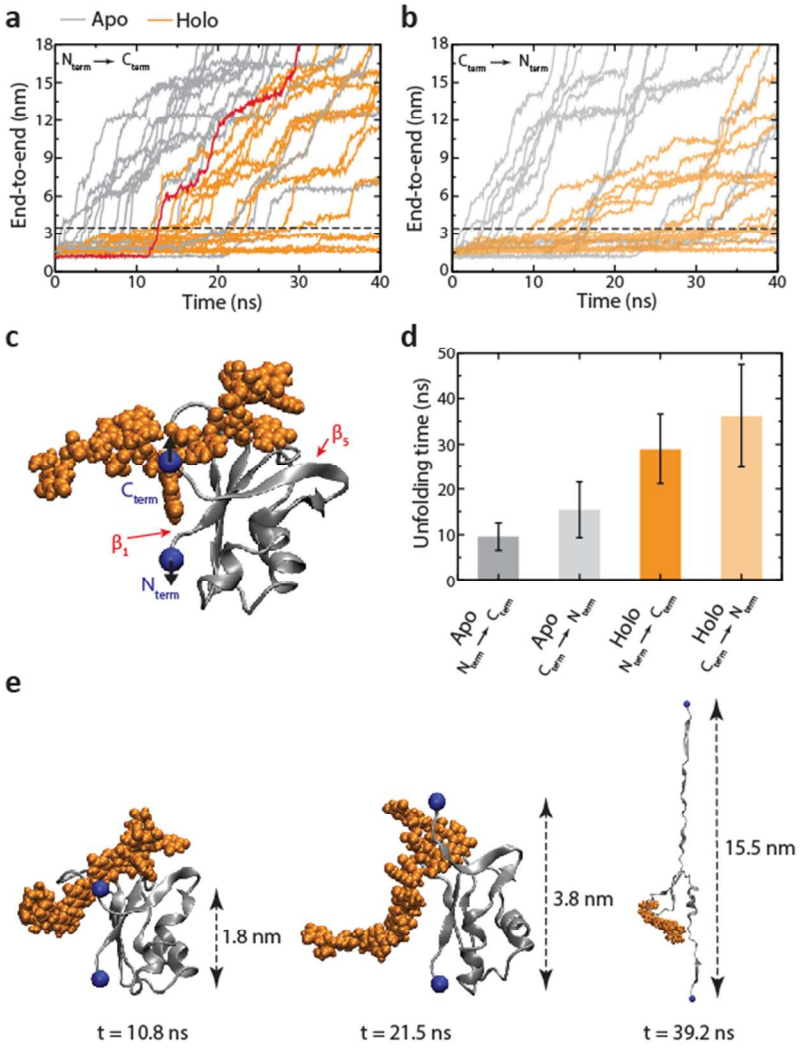


Figure 4. (a) End-to-end as a function of time for 15 simulations in the apo state (grey) and 15 simulations in the holo state (orange), under a constant force of 160 pN, and pulling on the C-terminal. Note that for trajectories where the protein did not unfold within the 40-ns time window shown here, the simulations were extended until unfolding occurred. One trajectory of the protein in the holo state is highlighted in red. (b) Same data but pulling on the N-terminal instead, which shows a slight increase in mechanical resistance for both apo- and holo- cases. (c) Solution NMR structure (PDB ID: 4BS2) of the RRM1 domain (grey) in complex with RNA (orange), where the residues interacting with RRM2 in the structure have been removed. The N- and C-terminal C_α, where force is applied (black arrows), are shown as blue balls. RRM1 mechanical unfolding is triggered by the loss of the interaction between ß1 and ß5. (d) Mean unfolding times, showing that the presence of RNA considerably increases the mechanical resistance of the RRM1 protein. Proteins are considered unfolded when their end-to-end distance exceeds 3.5 nm (dashed black line in panels (a) and (b)). (e) Snapshots of protein/RNA configurations at representative time intervals corresponding to an individual unfolding trajectory (red in (a): (*left*) stable initial intermediate before unfolding, showing that RNA acts as a mechanical "lid" that momentarily blocks unfolding; local fluctuations of the RNA-protein binding site allow detachment of ß5 from ß1 (*middle*), which triggers the subsequent unfolding of the protein (*right*).

About 2% of the human proteome consists of DNA- and RNA- binding proteins (DRBPs)[33]. As expected, DRBPs are functionally flexible, and are mainly involved in transcriptional regulation, mRNA processing and DNA replication. The versatile RRM (RNA recognition motif) domain is the most enriched domain in DRBPs, underpinning a structure able

to simultaneously bind single stranded ssRNA and ssDNA, as well as proteins. TDP-43 is a paradigmatic DRBP that has important roles in mRNA splicing[48] and miRNA biogenesis[52], and is formed by two independent RRM domains. Here we made use of multidisciplinary approach including biochemistry and molecular biology techniques combined with single molecule nanomechanical experiments complemented by SMD simulations to demonstrate that nucleic acid binding can result in the mechanical stabilization of the RRM1 domain of TDP-43.

The main discovery of our single molecule experiments is that, under the presence of a stretching force, the apo-form extends in a feature-less manner, demonstrating that the forces required to unfold are lower than ~ 15 pN, the intrinsic resolution of the AFM working under constant velocity conditions. Further work using e.g. optical or magnetic tweezers, with expanded low-force resultion[53], could provide further insight into the individual unfolding pathways of the apo-form of RRM1. By contrast, the addition of TG nucleotides results in a remarkable mechanical stabilization of RRM1, giving rise to the characteristic saw-tooth pattern of unfolding. However, the distribution of unfolding trajectories corresponding to the apo-form (and also to the control experiments in the presence of CA) shows that a finite number of unfolding events (~30 %) exhibiting mechanical resistance is always present. It is possible that after our purification, and despite quantitative DNA removal[46], there is still a low fraction of DNA-bound proteins. Conversely, in the presence of TG_{15}, we also observed a number of unfolding events (~17 %) devoid of mechanical stability. We attribute these trajectories to the presence of a dynamic binding/unbinding kinetics of the ssDNA oligonucleotides. Furthermore, our titration experiments (Figure S3) demonstrate that, as expected, the short TG_3 shows a lower binding affinity to RRM1 than TG_4, and that, in general, the binding affinity increases with the

length of the TG oligonucleotide. Notably, the different extent of binding of the TG oligonucleotides of different lengths nicely correlates with the fraction of trajectories showing mechanical stabilization (Figure 2c). Hence, an important conclusion stemming from our experiments is that mechanical stability appears to be a direct reporter of DNA binding.

To obtain insight into the molecular origin of the mechanical stabilization process, we conducted SMD simulations under force clamp conditions. Our simulations revealed a plausible mechanism whereby the oligonucleotide binds directly onto the mechanical clamp of RRM1 (ß1 and ß5) and acts as a 'mechanical lid' that locks the protein, preventing unfolding. This protective mechanism largely differs for example from the 'allosteric' binding mechanism described for protein G, whereby the IgG ligand was found to bind on a position away from the mechanical clamp[26]. Thermal fluctuations coupled with the effects of the pulling force are able to perturb the location of the nucleic acid, eventually slightly displacing it from the protein binding site and thus allowing the detachment of ß5 from ß1, which triggers the subsequent unfolding of the protein. Hence, given that mechanical stabilisation is mostly dictated by the residence time of the oligonucleotide bound to RRM1, it is tempting to speculate that in the presence of oligonucleotides able to bind RRM1 with high affinity (i.e. long TG repeats), the protein would take longer to unfold, thus giving rise to an apparent higher mechanical stability under force extension conditions. Indeed, close inspection to the histograms reporting the experimental unfolding forces for the different TG constructs (Figure S4 and Table S1) shows an overall significant ($p < 0.05$) increase of mechanical stability for those constructs with longer TG repeats, shifting from ~25 pN for TG_3 up to ~40 pN in the case of TG_{15}, thus qualitatively supporting the bound-stabilization argument. From the experimental perspective, it remains to be

seen whether RNA binding gives rise to a similar mechanical stabilization to that obtained by DNA-related sequences. A further interesting feature observed in our long simulations is that, after being displaced from the binding pocket enabling protein unfolding, the oligonucleotide is kept bound to the mechanically stretched protein (Supplementary video 1). It is hence possible that the relatively fast rebinding that we observed in our force-clamp experiments (Figure 3a) occurs because the oligonucleotide remains bound after mechanical unfolding, hence rendering the rebinding process more efficient.

From a broader perspective, our findings demonstrate that, akin to the mechanical stabilization effect triggered by protein-binding[12], nucleic acid binding can also have a drastic effect on the mechanical stability of proteins. Most significantly, our results show that, rather than changing the unfolding pathway as observed in the case of p53[34], DNA binding can directly affect the mechanical stability of a protein under force. Given the emerging role of the nucleus as a mechanosensor[54-55], and in light of recent discoveries evidencing that external mechanical perturbations can trigger gene expression after DNA remodelling[56-58], it is at least tantalising to hypothesize that nucleic acid binding can have important effects on the conformational dynamics of e.g transcription factors. In particular, it is plausible that such a conformational stabilization can play a key role in the dynamics of exon 9 skipping of the human *CFTR* gene, in the regulation of which TDP-43 plays a key role[48, 50]. Although almost all the ALS and FTD pathological mutations map to the intrinsically disordered C-terminal region and also to the RRM2[59] and N-terminal domain of TDP-43[60], the delayed unfolding upon DNA binding demonstrated here for RRM1 might be generally related to a delayed onset or inhibition of TDP-43 aggregation[49] –the hallmark of amyotrophic lateral sclerosis (ALS) and frontotemporal lobar

degeneration (FTLD). Moreover, it is also possible that the TDP-43 stabilization upon nucleic acid binding that we observe might be related to the increased solubilization of nuclear TDP-43[61]. Further single molecule experiments on the other regions of the TDP-43 protein, directed to elucidate how nucleic acid binding affects the nanomechanical properties of the RRM2, the C-terminal region and the NLS sequence[42], would increase the biological relevance of the reported findings. From a strict mechanical viewpoint, our results suggest a new rational way to modulate the mechanical properties of proteins to add to post-translational modifications[62-64] or ligand binding[12], and highlights the use of protein mechanical stability as an emerging molecular reporter of nucleic acid binding to proteins.

ASSOCIATED CONTENT

Supporting Information.

Supplementary material for this article is available at:

Materials and Methods

Figure S1. Distribution of the increment in contour length values associated to the unfolding of RRM.

Figure S2. Gallery of unfolding trajectories of the [$I27_2$-RRM1-I27-RRM1-$I27_2$] polyprotein.

Figure S3. Protein:DNA titrations monitored by EMSA.

Figure S4. Distribution of unfolding forces for those trajectories featuring mechanical stability**.**

Figure S5. Distributions of the step size (a) and force (b) associated to the unfolding of RRM1 in the TG_{15}-bound holo- form in the force ramp experiment.

Figure S6. Refolding trajectory of the [$I27_2$-RRM1-I27-RRM1-$I27_2$] polyprotein in the presence of TG_6.

Figure S7. Simulation of the end-to-end distance as a function of time for proteins with frozen RNA.

Table S1. Statistical analysis of the unfolding forces under the different experimental conditions.

Supplementary Movie 1. Individual SMD trajectory of RRM1 unfolding in the presence of RNA.

AUTHOR INFORMATION

The authors declare no competing commercial interests.

ACKNOWLEDGMENTS

We thank Dr. Amy Beedle for help in discussion of the results and critical reading of the manuscript. G.S. acknowledges support from CNRS through a PICS allocation (PICS07571) and from the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABX-0011-01). This work was supported by the BBSRC grant BB/J00992X/1, BHF grant (PG/13/50/30426), EPSRC Fellowship K00641X/1, by the European Commission (grant agreement SEP-210342844), and by the Leverhulme Trust Research Leadership Award (RL-2016-015), all to S.G-M.

REFERENCES

1. Vogel, V., Mechanotransduction involving multimodular proteins: converting force into biochemical signals. *Annu Rev Biophys Biomol Struct* **2006,** *35*, 459-88.
2. Roca-Cusachs, P.; Iskratsch, T.; Sheetz, M. P., Finding the weakest link: exploring integrin-mediated mechanical molecular pathways. *J Cell Sci* **2012,** *125*, 3025-38.
3. Li, H.; Linke, W. A.; Oberhauser, A. F.; Carrion-Vazquez, M.; Kerkvliet, J. G.; Lu, H.; Marszalek, P. E.; Fernandez, J. M., Reverse engineering of the giant muscle protein titin. *Nature* **2002,** *418*, 998-1002.
4. Aubin-Tam, M. E.; Olivares, A. O.; Sauer, R. T.; Baker, T. A.; Lang, M. J., Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. *Cell* **2011,** *145*, 257-67.
5. Maillard, R. A.; Chistol, G.; Sen, M.; Righini, M.; Tan, J.; Kaiser, C. M.; Hodges, C.; Martin, A.; Bustamante, C., ClpX(P) generates mechanical force to unfold and translocate its protein substrates. *Cell* **2011,** *145*, 459-69.
6. Stirnemann, G.; Giganti, D.; Fernandez, J. M.; Berne, B. J., Elasticity, structure, and relaxation of extended proteins under force. *Proc Natl Acad Sci U S A* **2013,** *110*, 3847-52.
7. Li, H.; Carrion-Vazquez, M.; Oberhauser, A. F.; Marszalek, P. E.; Fernandez, J. M., Point mutations alter the mechanical stability of immunoglobulin modules. *Nat Struct Biol* **2000,** *7*, 1117-20.

8. Carl, P.; Kwok, C. H.; Manderson, G.; Speicher, D. W.; Discher, D. E., Forced unfolding modulated by disulfide bonds in the Ig domains of a cell adhesion molecule. *Proc Natl Acad Sci U S A* **2001,** *98*, 1565-70.
9. Wiita, A. P.; Ainavarapu, S. R.; Huang, H. H.; Fernandez, J. M., Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques. *Proc Natl Acad Sci U S A* **2006,** *103*, 7222-7.
10. Zheng, P.; Li, H., Highly covalent ferric-thiolate bonds exhibit surprisingly low mechanical stability. *J Am Chem Soc* **2011,** *133*, 6791-8.
11. Beedle, A. E.; Lezamiz, A.; Stirnemann, G.; Garcia-Manyes, S., The mechanochemistry of copper reports on the directionality of unfolding in model cupredoxin proteins. *Nat Commun* **2015,** *6*, 7894.
12. Hu, X.; Li, H., Force spectroscopy studies on protein-ligand interactions: a single protein mechanics perspective. *FEBS Lett* **2014,** *588*, 3613-20.
13. Ainavarapu, S. R.; Li, L.; Badilla, C. L.; Fernandez, J. M., Ligand binding modulates the mechanical stability of dihydrofolate reductase. *Biophys J* **2005,** *89*, 3337-44.
14. Wang, C. C.; Tsong, T. Y.; Hsu, Y. H.; Marszalek, P. E., Inhibitor binding increases the mechanical stability of staphylococcal nuclease. *Biophys J* **2011,** *100*, 1094-9.
15. Bertz, M.; Rief, M., Ligand binding mechanics of maltose binding protein. *J Mol Biol* **2009,** *393*, 1097-105.
16. Aggarwal, V.; Kulothungan, S. R.; Balamurali, M. M.; Saranya, S. R.; Varadarajan, R.; Ainavarapu, S. R., Ligand-modulated parallel mechanical unfolding pathways of maltose-binding proteins. *J Biol Chem* **2011,** *286*, 28056-65.
17. Rivas-Pardo, J. A.; Alegre-Cebollada, J.; Ramirez-Sarmiento, C. A.; Fernandez, J. M.; Guixe, V., Identifying sequential substrate binding at the single-molecule level by enzyme mechanical stabilization. *ACS Nano* **2015,** *9*, 3996-4005.
18. Zocher, M.; Zhang, C.; Rasmussen, S. G.; Kobilka, B. K.; Muller, D. J., Cholesterol increases kinetic, energetic, and mechanical stability of the human beta2-adrenergic receptor. *Proc Natl Acad Sci U S A* **2012,** *109*, E3463-72.
19. Kotamarthi, H. C.; Sharma, R.; Narayan, S.; Ray, S.; Ainavarapu, S. R., Multiple unfolding pathways of leucine binding protein (LBP) probed by single-molecule force spectroscopy (SMFS). *J Am Chem Soc* **2013,** *135*, 14768-74.
20. Junker, J. P.; Rief, M., Single-molecule force spectroscopy distinguishes target binding modes of calmodulin. *Proc Natl Acad Sci U S A* **2009,** *106*, 14361-6.
21. Junker, J. P.; Ziegler, F.; Rief, M., Ligand-dependent equilibrium fluctuations of single calmodulin molecules. *Science* **2009,** *323*, 633-7.
22. Ramanujam, V.; Kotamarthi, H. C.; Ainavarapu, S. R., Ca2+ binding enhanced mechanical stability of an archaeal crystallin. *PLoS One* **2014,** *9*, e94513.
23. Oroz, J.; Valbuena, A.; Vera, A. M.; Mendieta, J.; Gomez-Puertas, P.; Carrion-Vazquez, M., Nanomechanics of the cadherin ectodomain: "canalization" by Ca2+ binding results in a new mechanical element. *J Biol Chem* **2011,** *286*, 9405-18.
24. Cao, Y.; Yoo, T.; Li, H., Single molecule force spectroscopy reveals engineered metal chelation is a general approach to enhance mechanical stability of proteins. *Proc Natl Acad Sci U S A* **2008,** *105*, 11152-7.
25. Cao, Y.; Li, Y. D.; Li, H., Enhancing the mechanical stability of proteins through a cocktail approach. *Biophys J* **2011,** *100*, 1794-9.

26. Cao, Y.; Li, H., Engineered elastomeric proteins with dual elasticity can be controlled by a molecular regulator. *Nat Nanotechnol* **2008,** *3*, 512-6.
27. Kotamarthi, H. C.; Yadav, A.; Koti Ainavarapu, S. R., Small peptide binding stiffens the ubiquitin-like protein SUMO1. *Biophys J* **2015,** *108*, 360-7.
28. Perales-Calvo, J.; Lezamiz, A.; Garcia-Manyes, S., The Mechanochemistry of a Structural Zinc Finger. *J Phys Chem Lett* **2015,** *6*, 3335-40.
29. Friedland, J. C.; Lee, M. H.; Boettiger, D., Mechanically activated integrin switch controls alpha5beta1 function. *Science* **2009,** *323*, 642-4.
30. del Rio, A.; Perez-Jimenez, R.; Liu, R.; Roca-Cusachs, P.; Fernandez, J. M.; Sheetz, M. P., Stretching single talin rod molecules activates vinculin binding. *Science* **2009,** *323*, 638-41.
31. Yao, M.; Qiu, W.; Liu, R.; Efremov, A. K.; Cong, P.; Seddiki, R.; Payre, M.; Lim, C. T.; Ladoux, B.; Mege, R. M.; Yan, J., Force-dependent conformational switch of alpha-catenin controls vinculin binding. *Nat Commun* **2014,** *5*, 4525.
32. Bechtluft, P.; van Leeuwen, R. G.; Tyreman, M.; Tomkiewicz, D.; Nouwen, N.; Tepper, H. L.; Driessen, A. J.; Tans, S. J., Direct observation of chaperone-induced changes in a protein folding pathway. *Science* **2007,** *318*, 1458-61.
33. Hudson, W. H.; Ortlund, E. A., The structure, function and evolution of proteins that bind DNA and RNA. *Nat Rev Mol Cell Biol* **2014,** *15*, 749-60.
34. Taniguchi, Y.; Kawakami, M., Variation in the mechanical unfolding pathway of p53DBD induced by interaction with p53 N-terminal region or DNA. *PLoS One* **2012,** *7*, e49003.
35. Wang, I. F.; Wu, L. S.; Shen, C. K., TDP-43: an emerging new player in neurodegenerative diseases. *Trends Mol Med* **2008,** *14*, 479-85.
36. Buratti, E.; Baralle, F. E., TDP-43: gumming up neurons through protein-protein and protein-RNA interactions. *Trends Biochem Sci* **2012,** *37*, 237-47.
37. Warraich, S. T.; Yang, S.; Nicholson, G. A.; Blair, I. P., TDP-43: a DNA and RNA binding protein with roles in neurodegenerative diseases. *Int J Biochem Cell Biol* **2010,** *42*, 1606-9.
38. Neumann, M.; Sampathu, D. M.; Kwong, L. K.; Truax, A. C.; Micsenyi, M. C.; Chou, T. T.; Bruce, J.; Schuck, T.; Grossman, M.; Clark, C. M.; McCluskey, L. F.; Miller, B. L.; Masliah, E.; Mackenzie, I. R.; Feldman, H.; Feiden, W.; Kretzschmar, H. A.; Trojanowski, J. Q.; Lee, V. M., Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **2006,** *314*, 130-3.
39. Banks, G. T.; Kuta, A.; Isaacs, A. M.; Fisher, E. M., TDP-43 is a culprit in human neurodegeneration, and not just an innocent bystander. *Mamm Genome* **2008,** *19* (5), 299-305.
40. Ederle, H.; Dormann, D., TDP-43 and FUS en route from the nucleus to the cytoplasm. *FEBS Lett* **2017,** *591*, 1489-1507.
41. Passoni, M.; De Conti, L.; Baralle, M.; Buratti, E., UG repeats/TDP-43 interactions near 5' splice sites exert unpredictable effects on splicing modulation. *J Mol Biol* **2012,** *415*, 46-60.
42. Mompean, M.; Romano, V.; Pantoja-Uceda, D.; Stuani, C.; Baralle, F. E.; Buratti, E.; Laurents, D. V., The TDP-43 N-terminal domain structure at high resolution. *FEBS J* **2016,** *283*, 1242-60.
43. Kuo, P. H.; Doudeva, L. G.; Wang, Y. T.; Shen, C. K.; Yuan, H. S., Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Res* **2009,** *37*, 1799-808.

44. Lukavsky, P. J.; Daujotyte, D.; Tollervey, J. R.; Ule, J.; Stuani, C.; Buratti, E.; Baralle, F. E.; Damberger, F. F.; Allain, F. H., Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat Struct Mol Biol* **2013,** *20*, 1443-9.
45. Kuo, P. H.; Chiang, C. H.; Wang, Y. T.; Doudeva, L. G.; Yuan, H. S., The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic Acids Res* **2014,** *42*, 4712-22.
46. Glasel, J. A., Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *Biotechniques* **1995,** *18*, 62-3.
47. Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M., Mechanical and chemical unfolding of a single protein: a comparison. *Proc Natl Acad Sci U S A* **1999,** *96*, 3694-9.
48. Buratti, E.; Dork, T.; Zuccato, E.; Pagani, F.; Romano, M.; Baralle, F. E., Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J* **2001,** *20*, 1774-84.
49. Huang, Y. C.; Lin, K. F.; He, R. Y.; Tu, P. H.; Koubek, J.; Hsu, Y. C.; Huang, J. J., Inhibition of TDP-43 aggregation by nucleic acid binding. *PLoS One* **2013,** *8*, e64002.
50. Buratti, E.; Baralle, F. E., Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *J Biol Chem* **2001,** *276*, 36337-43.
51. Fernandez, J. M.; Li, H., Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science* **2004,** *303*, 1674-8.
52. Kawahara, Y.; Mieda-Sato, A., TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc Natl Acad Sci U S A* **2012,** *109*, 3347-52.
53. Rivas-Pardo, J. A.; Eckels, E. C.; Popa, I.; Kosuri, P.; Linke, W. A.; Fernandez, J. M., Work Done by Titin Protein Folding Assists Muscle Contraction. *Cell Rep* **2016,** *14*, 1339-1347.
54. Fedorchak, G. R.; Kaminski, A.; Lammerding, J., Cellular mechanosensing: getting to the nucleus of it all. *Prog Biophys Mol Biol* **2014,** *115*, 76-92.
55. Cho, S.; Irianto, J.; Discher, D. E., Mechanosensing by the nucleus: From pathways to scaling relationships. *J Cell Biol* **2017,** *216*, 305-315.
56. Tajik, A.; Zhang, Y.; Wei, F.; Sun, J.; Jia, Q.; Zhou, W.; Singh, R.; Khanna, N.; Belmont, A. S.; Wang, N., Transcription upregulation via force-induced direct stretching of chromatin. *Nat Mater* **2016,** *15*, 1287-1296.
57. Poh, Y. C.; Shevtsov, S. P.; Chowdhury, F.; Wu, D. C.; Na, S.; Dundr, M.; Wang, N., Dynamic force-induced direct dissociation of protein complexes in a nuclear body in living cells. *Nat Commun* **2012,** *3*, 866.
58. Jain, N.; Iyer, K. V.; Kumar, A.; Shivashankar, G. V., Cell geometric constraints induce modular gene-expression patterns via redistribution of HDAC3 regulated by actomyosin contractility. *Proc Natl Acad Sci U S A* **2013,** *110*, 11349-54.
59. Guenther, E. L.; Ge, P.; Trinh, H.; Sawaya, M. R.; Cascio, D.; Boyer, D. R.; Gonen, T.; Zhou, Z. H.; Eisenberg, D. S., Atomic-level evidence for packing and positional amyloid polymorphism by segment from TDP-43 RRM2. *Nat Struct Mol Biol* **2018,** *25*, 311-319.
60. Wang, A.; Conicella, A. E.; Schmidt, H. B.; Martin, E. W.; Rhoads, S. N.; Reeb, A. N.; Nourse, A.; Ramirez Montero, D.; Ryan, V. H.; Rohatgi, R.; Shewmaker, F.; Naik, M. T.; Mittag, T.; Ayala, Y. M.; Fawzi, N. L., A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *EMBO J* **2018,** *37*, e97452.

61. Maharana, S.; Wang, J.; Papadopoulos, D. K.; Richter, D.; Pozniakovsky, A.; Poser, I.; Bickle, M.; Rizk, S.; Guillen-Boixet, J.; Franzmann, T. M.; Jahnel, M.; Marrone, L.; Chang, Y. T.; Sterneckert, J.; Tomancak, P.; Hyman, A. A.; Alberti, S., RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **2018,** *360*, 918-921.
62. Alegre-Cebollada, J.; Kosuri, P.; Giganti, D.; Eckels, E.; Rivas-Pardo, J. A.; Hamdani, N.; Warren, C. M.; Solaro, R. J.; Linke, W. A.; Fernandez, J. M., S-glutathionylation of cryptic cysteines enhances titin elasticity by blocking protein folding. *Cell* **2014,** *156*, 1235-1246.
63. Beedle, A. E.; Lynham, S.; Garcia-Manyes, S., Protein S-sulfenylation is a fleeting molecular switch that regulates non-enzymatic oxidative folding. *Nat Commun* **2016,** *7*, 12490.
64. Beedle, A. E. M.; Mora, M.; Lynham, S.; Stirnemann, G.; Garcia-Manyes, S., Tailoring protein nanomechanics with chemical reactivity. *Nat Commun* **2017,** *8*, 15658.